Journal of Cardiovascular Nursing
Vol. 00, No. 0, pp 00–00 | Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

Psychometric Properties of the 9-item European Heart Failure Self-Care Behavior Scale Using Confirmatory Factor Analysis and Rasch Analysis Among Iranian Patients

Chung-Ying Lin, PhD; Amir H. Pakpour, PhD; Anders Broström, PhD; Bengt Fridlund, PhD; Kristofer Årestedt, PhD; Anna Strömberg, PhD; Tiny Jaarsma, PhD; Jan Mårtensson, PhD

Background: The 9-item European Heart Failure Self-Care Behavior scale (EHFScB-9) is a self-reported questionnaire commonly used to capture the self-care behavior of people with heart failure (HF). **Objective:** The aim of this study was to investigate the EHFScB-9's factorial structure and categorical functioning of the response scale and differential item functioning (DIF) across subpopulations in Iran. **Methods:** Patients with HF (n = 380; 60.5% male; mean [SD] age, 61.7 [9.1] years) participated in this study. The median (interquartile range) of the duration of their HF was 6.0 (2.4–8.8) months. Most of the participants were in New York Heart Association classification II (NYHA II, 61.8%); few of them had left ventricular ejection fraction assessment (11.3%). All participants completed the EHFScB-9. Confirmatory factor analysis was used to test the factorial structure of the EHFScB-9; Rasch analysis was used to analyze categorical functioning and DIF items across 2 characteristics (gender and NYHA). **Results:** The 2-factor structure (''adherence to regimen'' and ''consulting behavior'') of the EHFScB-9 was confirmed, and the unidimensionality of each factor was found. Categorical functioning was supported for all items. No items displayed substantial DIF across gender (DIF contrast, −0.25–0.31). Except for item 3 (''Contact doctor or nurse if legs/feet are swollen''; DIF contrast, −0.69), no items displayed substantial DIF across NYHA classes (DIF contrast, −0.40 to 0.47). **Conclusions:** Despite the DIF displayed in 1 item across the NYHA classes, the EHFScB-9 demonstrated sound psychometric properties in patients with HF.

KEY WORDS: confirmatory factor analysis, heart failure, Rasch, self-care behavior scale

Heart failure (HF)—a complex syndrome with high rates of mortality and morbidity[1,2]—remains a major health problem globally.[1–4] Recent data from the American Heart Association show a significantly increased number: more than 6.5 million people in the United States and more than 15 million in Europe experience HF.[1] The prevalence of HF in developed countries is approximately 1% to 2% and much higher (>10%) if only focusing on older populations older than 70 years.[5,6] The prevalence of HF in developing countries is high as well; for example, the prevalence in Iran in the near future is projected to be approximately 3.5%.[7]

To decrease the social and caregiver burdens from HF, the European Society of Cardiology guidelines for acute and chronic HF propose that patients with HF should receive a holistic and multidisciplinary approach,

Chung-Ying Lin, PhD
Assitatnt Professor, Faculty of Health and Social Sciences, Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hung Hom.
Amir H. Pakpour, PhD
Associate Professor, Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Iran, and Department of Nursing, School of Health and Welfare, Jönköping University, Sweden.
Anders Broström, PhD
Professor, Department of Nursing, School of Health and Welfare, Jönköping University, Sweden.
Bengt Fridlund, PhD
Professor, School of Health and Welfare, Jönköping University, Sweden.
Kristofer Årestedt, PhD
Professor, Faculty of Health and Life Sciences, Linnaeus University, Kalmar, and Department of Medical and Health Sciences, Linköping University, Sweden.
Anna Strömberg, PhD
Professor, Department of Medical and Health Sciences, Linköping University, Sweden.
Tiny Jaarsma, PhD
Professor, Department of Social and Welfare Studies, Linköping University, Sweden; and Mary MacKillop Institute of Health Research, ACU, Melbourne, Australia.
Jan Mårtensson, PhD
Professor, Department of Nursing, School of Health and Welfare, Jönköping University, Sweden.

The authors have no funding or conflicts of interest to disclose.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jcnjournal.com).

Correspondence
Amir H. Pakpour, PhD, Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Shahid Bahonar Blvd, Qazvin 3419759811, Iran (pakpour_amir@yahoo.com).

DOI: 10.1097/JCN.0000000000000444

including psychoeducation and promotion of self-care behavior.[5,6] Self-care behavior for HF is defined as "a naturalistic decision-making process involving the choice of behaviors that maintain physiologic stability (maintenance) and the response to symptoms when they occur (management)."[8,9] Patients with HF can improve their health if they engage in such effective behaviors.[10,11] Therefore, given that the importance of self-care behavior for patients with HF is highlighted, an effective and feasible instrument measuring such behaviors is warranted.

The most commonly used instruments developed for self-care behavior for patients with HF are the Self-Care of Heart Failure Index[12] and the European Heart Failure Self-Care Behavior scale (EHFScB).[13] However, Self-Care of Heart Failure Index and EHFScB have different conceptualizations of HF self-care: Self-Care of Heart Failure Index focuses on 3 dimensions of self-care (maintenance, management, and confidence); EHFScB was developed first using the 3 dimensions of complying with regimen, asking for help, and adapting activities. We investigated the use of EHFScB because it has the following strengths: (1) short and easy to understand, (2) developed through a strong procedure by many HF experts,[13] (3) useful to evaluate the effectiveness of psychoeducation on self-care behavior because the EHFScB covers the patient's ability to participate in effective self-care, and (4) a theory-driven approach to design a 2-factor structure (ie, adherence to regimen and consulting behavior) instead of creating the factorial structure based on data properties. The EHFScB originally contained 12 items but has been shortened to 9 items (EHFScB-9) because of improvements in psychometric properties and reduced respondent burden.[14]

Although the EHFScB-9 is a promising instrument to assess HF self-care behavior, its psychometric properties are underdeveloped. Several studies have reported on the 12-item version, but only 5 studies have been identified that tested the psychometric properties of the EHFScB-9.[15] In addition, studies found different factorial structures for the EHFScB-9 using confirmatory factor analysis (CFA).[15] There are 4 proposed factorial structures for the EHFScB-9: an overall underlying domain[16]; 2 domains, "adherence to regimen" and "consulting behavior"[14]; 3 domains, "adhering to recommendations," "fluid and sodium management," and "physical activity and recognition of deteriorating symptoms"[17]; and 3 domains, "consulting behavior," "autonomy-based adherence," and "provider-based adherence."[18] Inconsistent findings regarding the factorial structure of the scale give no definite answers about the underlying structure of the scale; therefore, additional analyses comparing different factorial structures are needed. Østergaard et al[19] identified the problems of different factorial structures and used several CFAs to compare the 4 proposed factorial structures. They found that the 2-factor structure ("adherence to regimen" and "consulting behavior") outperformed other factorial structures; however, further corroborations on their findings are needed. Specifically, they did not use diagonally weighted least squares, an estimator for Likert-type scales,[20,21] to take care of their models; they treated the indicator variables (ie, items) as continuous, and the estimated parameters (eg, loading and standard error) may be biased.

Another underdeveloped issue for the psychometric properties of the EHFScB-9 regards the theory used for psychometric testing. Studies have mainly adopted the classical test theory to evaluate the psychometric properties of the EHFScB-9.[14,16–19] Although the internal consistency of the EHFScB-9 was satisfactory across several European countries using Cronbach's α,[14,19] to the best of our knowledge, no one has used Rasch models. Rasch analysis uses a mathematical formula to estimate the probability of an individual responding with a certain answer (eg, scores 1–5 in the Likert-type scale) on an item. Using the formula, both item and person properties can be estimated; that is, we can estimate the item difficulty (whether the HF self-care is easy to be complied with) and the person's ability (whether the person has the capability to do the HF self-care). We acknowledge that using classical test theory has the benefit of being easily understood[22]; however, without the results analyzed using Rasch models, some psychometric characteristics of the EHFScB-9 cannot be identified. For example, it is unclear how each item is embedded in the EHFScB-9 (ie, how much the concept of the item is out of the EHFScB-9 or how much of it is redundant to the information provided by other items), whether the categorical functioning of the response scale is in order, and whether different subpopulations of patients with HF interpret the EHFScB-9 items similarly.

Testing whether different subgroups interpret the EHFScB-9 items differently is a critical issue because the score differences of a measure can be the true group differences or the various understandings toward the same measure contents.[23] If we want to ensure that the score difference in the EHFScB-9 is from the true group difference, we should exclude the possibility of various understandings. Differential item functioning (DIF) in the Rasch analysis is a recommended method to tackle this issue.[24,25]

This study aimed to (1) compare competing CFA models previously described in the literature and (2) evaluate the instrument using the Rasch model.

Methods

Instrument

The EHFScB-9 contains 9 items measuring how well a patient with HF manages self-care. The item contents

regarding how to self-care were designed specifically on HF symptoms, and all the items are rated on a 5-point Likert scale ranging from 1 (completely agree) to 5 (completely disagree). A lower score represents better self-care of a patient with HF.[14]

Design

The ethics committee of Qazvin University of Medical Sciences approved the study. All participants gave written informed consent for inclusion in the study. The psychometric evaluation study was conducted in 2 steps: (1) translation into Persian and pilot-testing the EHFScB-9[26,27] and (2) psychometric evaluation. Detailed information of the first step is described in Appendix A, Supplemental Digital Content 1, http://links.lww.com/JCN/A42.

Participants and Procedures for Psychometric Evaluation

Patients with HF were recruited from 7 university hospitals of Tehran, Qazvin, Tabriz, Mashhad, Zanjan, Sari, and Ilam. All hospitals were referral university hospitals that have the highest admission rates of heart diseases. Inclusion criteria were as follows: patients who are at least 18 years old, had New York Heart Association classification II to IV, and are able to read and write in Persian. Exclusion criteria were as follows: have cognitive impairment (Mini-Mental State Examination score < 25), received heart transplant, or have inability to understand Persian. Clinical variables were collected from patients' medical records.

Statistical Analysis

We tested the internal consistency of the EHFScB-9 using polychoric correlation (ρ)[28]: the ordinal α ($\alpha > .7$ suggests acceptable) and corrected item-total polychoric correlations ($\rho > 0.4$ suggests acceptable).

We used CFA to determine the factorial structure of the EHFScB-9. Specifically, we tested 4 different competing factorial models (Table 1) suggested in previous research: 1-factor model,[16] 2-factor model ("adherence to regimen" and "consulting behavior"),[14] and 3-factor model ("adherence to recommendations," "fluid and sodium management," and "physical activity and recognition of deteriorating symptoms") proposed by Lambrinou et al[17] and another 3-factor model ("consulting behavior," "autonomy-based adherence," and "provider-based adherence") proposed by Vellone et al.[18] We adopted the estimator of diagonally weighted least squares, which is based on polychoric correlations and recommended for Likert-type scales.[20,21] To identify supported models, we used comparative fit index and Tucker-Lewis index greater than 0.95, root mean square of error approximation less than 0.05, and weighted root mean square residual less than 0.90.[29–31] Moreover, we detected whether any model contained offending estimates (eg, correlation coefficient > 1) and considered the model with offending estimates as an inappropriate model (ie, Heywood cases).

After the factorial structure was confirmed by CFA, we used the Rasch partial credit model to examine the unidimensionality of each EHFScB-9 factor, item difficulty, item fit statistics, separation reliability, categorical functioning, and DIF for gender and New York Heart Association classification. We used principal component analysis on the standardized residuals retrieved from Rasch models to test the unidimensionality and the first component's eigenvalue of less than 2 to indicate unidimensionality.[22] Item difficulty was presented using the log-odds unit, namely, *logit*, where a higher logit represents a more difficult item. Item fit statistics included information-weighted fit statistic (infit) mean square (MnSq) and outlier-sensitive fit statistic (outfit) MnSq (the ratio of the observed response to the predicted response with an ideal value at 1.0); however, some variation from expectation (ie, 40%) is allowed because Rasch analysis is a probability model. Infit and outfit MnSqs between 0.6 and 1.4 of each item suggest an acceptable fit.[32] Separation reliability included person and item separation reliabilities; a value greater than 0.7 is acceptable.[33] Categorical functioning was used to understand whether successive response categories for each item were located in their expected order; for example, the difficulty of the response "completely agree" should be lower than that of the response "agree" in all items. Both average measure (the estimated average ability on a particular category) and step measure (the thresholds between categories) should monotonically increase with categories.[34] Differential item functioning was examined for each item across gender (male vs female) and across New York Heart Association classification (NYHA II vs NYHA III and IV), and a DIF contrast (ie, the logit of group 1 minus the logit of group 2) greater than 0.5 logit suggests substantial DIF.[35]

Descriptive statistics were analyzed using IBM SPSS 23.0 (IBM Corp, Armonk, New York), CFA and ordinal α using the R software (lavaan package for CFA[36] and psych for ordinal α[37]), and Rasch models using WINSTEPS software (Winsteps, Chicago, Illinois).

Results

The response rate was 87% (n = 380). Table 2 presents the demographics of participants. In brief, the participants were relatively young (mean ± SD age, 61.7 ± 9.1 years) with limited education (mean ± SD, 6.4 ± 3.7 years), and less than half of the participants were receiving guideline-based care (eg, ACE inhibitors [42.6%] and

TABLE 1 Model Comparisons of the 9-item European Heart Failure Self-Care Behavior Scale Using Confirmatory Factor Analysis (N = 380)

		Model 2		Model 3			Model 4		
	Model 1	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Factor loadings									
Item 1: weigh everyday	0.70	0.80			0.78			0.80	
Item 2: contact doctor or nurse if SOB is increased	0.82		0.85			0.80	0.85		
Item 3: contact doctor or nurse if legs/feet are swollen	0.68		0.71	0.66			0.71		
Item 4: contact doctor or nurse if gaining weight	0.82		0.85			0.80	0.85		
Item 5: limit amount of fluids	0.60	0.68			0.65			0.68	
Item 6: contact doctor or nurse if experiencing fatigue	0.80		0.83	0.78			0.83		
Item 7: eat low-salt diet	0.58	0.65			0.63				0.64
Item 8: take prescribed medication	0.60	0.67		0.59					0.66
Item 9: exercise regularly	0.55	0.62				0.54		0.62	
Factor correlations									
Factors 1 and 2		0.68			0.86			0.67	
Factors 1 and 3					1.10[a]			0.70	
Factors 2 and 3					0.88			1.02[a]	
Fit statistics									
χ^2 (*df*)	197.07 (27)	50.66 (26)		173.13 (24)			50.49 (24)		
P	<.001	.003		<.001			.001		
CFI	0.97	0.995		0.97			0.995		
TLI	0.96	0.993		0.96			0.992		
RMSEA	0.13	0.05		0.13			0.05		
90% CI of RMSEA	0.11–0.15	0.03–0.07		0.11–0.15			0.03–0.08		
WRMR	1.69	0.86		1.58			0.86		

Model 1 is proposed by Lee et al,[16] with only 1 factor.
Model 2 is proposed by Jaarsma et al,[14] with 2 factors: factor 1, adherence to regimen, and factor 2, consulting behavior.
Model 3 is proposed by Lambrinou et al,[17] with 3 factors: factor 1, adhering to recommendations; factor 2, fluid and sodium management; and factor 3, physical activity and recognition of deteriorating symptoms.
Model 4 is proposed by Vellone et al,[18] with 3 factors: factor 1, consulting behavior; factor 2, autonomy-based adherence; and factor 3, provider-based adherence.
Abbreviations: CFI, comparative fit index (>0.95 indicates good fit); TLI, Tucker-Lewis index (>0.95 indicates good fit); RMSEA, root mean square error of approximation (<0.05 indicates good fit); WRMR, weight root mean square residual (<0.90 indicates good fit); SOB, shortness of breath.
[a]The correlation coefficients are offending estimates. Therefore, both 3-factor models were unidentified.

TABLE 2 Participants Characteristics (N = 380)

	Mean ± SD	Median (IQR)	n (%)
Demographic variables			
Age, Years	61.7 ± 9.1	62.0 (55.0–69.0)	
Gender, male			230 (60.5)
Years of education	6.4 ± 3.7	5.0 (5.0–8.0)	
Body mass index, kg/m^2	28.2 ± 5.0	27.7 (24.7–31.5)	
Heart failure characteristics			
Duration of heart failure, Months	5.6 ± 3.7	6.0 (2.4–8.8)	
NYHA classification II			235 (61.8)
NYHA classification III			91 (23.9)
NYHA classification IV			54 (14.2)
Left ventricular ejection fraction, %			28.1 (11.3)
Medication, n (%)			
Diuretic			183 (48.2)
β-blockers			154 (40.5)
ACE inhibitors			162 (42.6)
Nitrate			47 (12.4)
Digoxin			50 (13.2)

Abbreviations: IQR, interquartile range; NYHA, New York Heart Association.

-blockers [40.5%]). The internal consistency was good for the entire EHFScB-9 (ordinal α = .87) and the 2 factors of the originally proposed 2-factor EHFScB-9 (ordinal α = .81 ["adherence to regimen"] and .89 ["consulting behavior"]). All the corrected item-total polyserial correlations were acceptable with the range between 0.46 and 0.67.

The originally proposed 2-factor model demonstrated better model fit than the 1- and 3-factor models suggested by Lambrinou et al.[17] The 3-factor model described by Vellone et al[18] showed all fit indices close to the 2-factor model; however, it contained a Heywood case: the correlation between factors 2 and 3 was larger than 1 (Table 1). On the basis of these results, the following Rasch analyses and concurrent validity were analyzed using the 2-factor structure: items 1, 5, 7, 8, and 9 were embedded in factor 1 ("adherence to regimen"), and items 2, 3, 4, and 6 were embedded in factor 2 ("consulting behavior").

Then, we conducted 2 Rasch models, one for each EHFScB-9 factor. The unidimensionality of each factor was supported by the first principal component analysis of the residuals: the first component's eigenvalue was 1.6 for "adherence to regimen" and 1.5 for "consulting behavior." All items fit well in their underlying construct. Although the range of item difficulties was slightly narrow, the difficulty ranges were widely distributed according to the item thresholds (ranged between −2.16 and 2.42 for "adherence to regimen" and between −4.33 and 4.04 for "consulting behavior"). Person separation reliability was satisfactory; item separation reliability was excellent (Table 3). Moreover, most of the item residuals showed low local independency (see Appendix B, Supplemental Digital Content 2, http://links.lww.com/JCN/A43, for details). The categorical functioning of each item was supported by average and step measures: both monotonically increased with categories (see Appendix C, Supplemental Digital Content 3, http://links.lww.com/JCN/A44, for details).

No items displayed substantial DIF across gender. Except for item 3 ("Contact doctor or nurse if legs/feet are swollen"), no items displayed substantial DIF across New York Heart Association classification. As for item 3, the DIF contrast (−0.69) suggested that the item description was easier for the participants with New York Heart Association classification II than those with New York Heart Association classification III or IV (Table 3). That is, participants with New York Heart Association classification II tended to score item 3 lower than those with New York Heart Association classification III or IV.

Discussion

Generally speaking, our results corroborate the findings of Østergaard et al[19] and Jaarsma et al[14]: the 2-factor structure model fits best for the EHFScB-9. Our excellent person and item separation reliabilities for both factors further corroborate the results of the 2-factor structure. The internal consistency found in our Iranian sample (α = .84) is satisfactory and consistent with those in the samples of European countries (α = .68–.87).[14,19] Our results additionally demonstrated the promising psychometric properties of the EHFScB-9. First, Rasch analyses supported the unidimensionality

TABLE 3 Rasch analysis (partial credit model) for the two factors of the 9-item European Heart Failure Self-Care Behavior Scale (N = 380)

		MnSq		Separation reliability		DIF contrast	
	Difficulty	Infit	Outfit	Person	Item	Gender	NYHA classification
Factor 1: Adherence to regimen				0.73	0.98		
Item 1: Weigh everyday	0.49	0.82	0.87			0.13	0.25
Item 5: Limit amount of fluids	0.49	0.98	1.01			0.00	−0.03
Item 7: Eat low-salt diet	−0.48	1.12	1.11			0.15	−0.40
Item 8: Take prescribed medication	−0.56	0.99	0.98			−0.21	0.34
Item 9: Exercise regularly	0.06	1.00	0.97			−0.06	−0.23
Factor 2: Consulting behavior				0.74	0.95		
Item 2: Contact doctor or nurse if SOB is increased	0.44	0.81	0.86			0.10	0.13
Item 3: Contact doctor or nurse if legs/feet are swollen	−0.78	1.25	1.32			−0.25	−0.69
Item 4: Contact doctor or nurse if gaining weight	0.23	0.94	0.95			−0.16	0.47
Item 6: Contact doctor or nurse if experiencing fatigue	0.11	0.96	0.97			0.31	0.14

DIF contrast for gender: difficulty in women – difficulty in men; DIF contrast for NYHA classification: difficulty in class II – difficulty in classes III and IV.
MnSq = mean square, values between 0.6 and 1.4 indicate good fit.
Separation reliability > 0.7 indicate acceptable reliability.
DIF = differential item functioning, value >1.5 indicates substantial DIF.
SOB = shortness of breath.
NYHA: New York Heart Association.

of the EHFScB-9 for each factor based on the 2-factor structure. Second, each item was well embedded in its underlying construct. Third, the categorical functioning of each item was in order followed by the difficulty, which suggests that the respondents could well distinguish the levels between scores 1 and 2, 2 and 3, and so on. Fourth, no items displayed substantial DIF across gender, and only item 3 ("Contact doctor or nurse if legs/feet are swollen") displayed substantial DIF across the classes of New York Heart Association classification (II vs III and IV).

Different factorial structures have been proposed for the EHFScB-9.[14,16–18] Østergaard et al[19] evaluated all the existing factorial structures using CFA and found that all the factorial structures did not have all fit indices being satisfactory, although the most favorable factorial structure was the 2-factor structure. Somewhat unsatisfactory fit indices were found in other studies across all of the proposed factorial structures[14,16–18]; however, we considered that the unsatisfactory fit indices could be attributable to the estimators used in the studies: except for Lee et al,[16] other studies[14,17–19] did not apply the diagonally weighted least squares estimator to account for the ordinal data and ordinal indicator variables in the EHFScB-9 (ie, the Likert-type scale). Although Lee et al[16] used the diagonally weighted least squares estimator, they only tested the 1-factor structure with somewhat promising fit indices (root mean square error of approximation, 0.12; weighted root mean square residual, 1.00; CFI, 0.94; TLI, 0.92), which were in accord with the fit indices of our 1-factor structure (root mean square error of approximation, 0.13; weighted root mean square residual, 1.69; comparative fit index, 0.97; Tucker-Lewis index, 0.96). Consequently, we adopted the diagonally weighted least squares estimator to retest all the proposed factorial structures to corroborate the findings of Østergaard et al,[19] and we strongly recommend using the 2-factor structure of the EHFScB-9 in future studies. However, cross-culture validation is warranted for further corroboration, that is, to investigate whether the 2 subscales are invariant measures of self-care behavior across different language versions.

In addition to using an appropriate estimator to reconfirm the factorial structure for EHFScB-9, another important finding in our study was the DIF items. Gender and severity levels are commonly used factors to test DIF. Taking the EHFScB-9 item "Contact doctor or nurse if gaining weight (more than 2 kg)" as an example, the difference in the item score could be because different genders (or different severity levels) have different self-care behaviors on this item. However, there is another possibility: men and women (or different severity levels) may have different perceptions of "gaining weight." If men (or patients with less severe HF) think that gaining 2.4 kg is not more than 2 kg because 0.4 kg may be due to error and indicates no need to contact doctor or nurse and women (or patients with more severe HF) think that gaining 2.01 kg indicates the need to contact doctor or nurse, comparing the item scores between genders (or different severity levels) would be inappropriate. Therefore, before using the EHFScB-9 score for comparisons, we need to investigate whether genders (or different severity levels) interpret the items similarly. Because most of the items did not display substantial DIF, we recommend using the EHFScB-9 items to compare self-care behavior between men and women or between those with New York Heart Association classification II and those with New York Heart Association classification III and IV.

However, when a researcher wants to compare self-care behavior across different New York Heart Association classification classes, attention should be paid to item 3 ("Contact doctor or nurse if legs/feet are swollen"). Our Rasch results indicate that participants with New York Heart Association classification II may more easily detect their legs/feet to become swollen than those with New York Heart Association classification III or IV. Because patients with New York Heart Association classification III and IV more frequently have severe symptoms than those with New York Heart Association classification II,[5] we postulated that patients with New York Heart Association classification III and IV are more used to symptoms and therefore may have more difficulties detecting changes in their swollen legs/feet. In contrast, patients with New York Heart Association classification II—who more seldom perceive these symptoms—may more easily observe such changes. Two limitations in our DIF results, however, should be noted. First, our analyses did not allow us to test whether DIF was real or artificial. It is important to identify DIF as real or artificial because it may violate the conclusion of invariant properties of an instrument.[38] Second, nonuniform DIF was not possible to evaluate. Future studies are needed to evaluate the importance of this DIF and whether it will affect the interpretation of scores between groups of different genders and New York Heart Association classification classes.

There are limitations in the study. First, we did not collect the retest data for the EHFScB-9; therefore, we cannot evaluate the reproducibility (test-retest reliability). The responsiveness and minimal clinically important differences are also left unknown because we did not collect such data. Second, we did not recruit patients with New York Heart Association classification I; thus, the generalizability of our CFA results may be limited.

In conclusion, despite 1 item that displayed DIF across the New York Heart Association classification classes, the EHFScB-9 demonstrated sound psychometric properties. The 2-factor structure ("adherence to regimen" and "consulting behavior") of the EHFScB-9

What's New and Important

- The 2-factor structure model (factors are ''adherence to regimen'' and ''consulting behavior'') fits the best for the EHFScB-9 in our advanced statistical methods (ie, the CFA using diagonally weighted least squares).
- Except for 1 item that displays DIF across New York Heart Association classification, no items displayed substantial DIF across gender and New York Heart Association classification.

was confirmed and supported by both CFA and Rasch analysis. Because the psychometric properties of the Persian EHFScB-9 were supported, our results could be generalizable to other Persian-speaking countries (Iran, Afghanistan, and Tajikistan, with ~120 million Persian-speaking people). A substantial amount of people may thus get benefits.

REFERENCES

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *Circulation*. 2017;135(10):e146–e603.
2. Grady KL. Self-care and quality of life outcomes in heart failure patients. *J Cardiovasc Nurs*. 2008;23(3):285–292.
3. Clark AM, Savard LA, Spaling MA, Heath S, Duncan AS, Spiers JA. Understanding help-seeking decisions in people with heart failure: a qualitative systematic review. *Int J Nurs Stud*. 2012;49(12):1582–1597.
4. Hwang B, Luttik ML, Dracup K, Jaarsma T. Family caregiving for patients with heart failure: types of care provided and gender differences. *J Card Fail*. 2010;16(5):398–403.
5. McMurray JJ, Adamopoulos S, Anker SD, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2012;33(14):1787–1847.
6. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37(27):2129–2200.
7. Bahrami M, Etemadifar S, Shahriari M, Farsani AK. Caregiver burden among Iranian heart failure family caregivers: a descriptive, exploratory, qualitative study. *Iran J Nurs Midwifery Res*. 2014;19(1):56–63.
8. Riegel B, Jaarsma T, Strömberg A. A middle-range theory of self-care of chronic illness. *ANS Adv Nurs Sci*. 2012;35(3):194–204.
9. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. *J Cardiovasc Nurs*. 2009;24(6):485–497.
10. Buck HG, Lee CS, Moser DK, et al. Relationship between self-care and health-related quality of life in older adults with moderate to advanced heart failure. *J Cardiovasc Nurs*. 2012;27(1):8–15.
11. Lee CS, Moser DK, Lennie TA, Tkacs NC, Margulies KB, Riegel B. Biomarkers of myocardial stress and systemic inflammation in patients who engage in heart failure self-care management. *J Cardiovasc Nurs*. 2011;26(4):321–328.
12. Vellone E, Riegel B, Cocchieri A, et al. Psychometric testing of the self-care of heart failure index version 6.2. *Res Nurs Health*. 2013;36(5):500–511.
13. Jaarsma T, Strömberg A, Mårtensson J, Dracup K. Development and testing of the European Heart Failure Self-Care Behaviour Scale. *Eur J Heart Fail*. 2003;5(3):363–370.
14. Jaarsma T, Arestedt KF, Mårtensson J, Dracup K, Strömberg A. The European Heart Failure Self-care Behaviour scale revised into a nine-item scale (EHFScB-9): a reliable and valid international instrument. *Eur J Heart Fail*. 2009;11(1):99–105.
15. Sedlar N, Socan G, Farkas J, et al. Measuring self-care in patients with heart failure: a review of the psychometric properties of the European Heart Failure Self-Care Behaviour Scale (EHFScBS). *Patient Educ Couns*. 2017;100(7):1304–1313.
16. Lee CS, Lyons KS, Gelow JM, et al. Validity and reliability of the European Heart Failure Self-care Behavior Scale among adults from the United States with symptomatic heart failure. *Eur J Cardiovasc Nurs*. 2013;12(2):214–218.
17. Lambrinou E, Kalogirou F, Lamnisos D, et al. The Greek version of the 9-item European Heart Failure Self-care Behaviour Scale: a multidimensional or a uni-dimensional scale? *Heart Lung*. 2014;43(6):494–499.
18. Vellone E, Jaarsma T, Strömberg A, et al. The European Heart Failure Self-care Behaviour Scale: new insights into factorial structure, reliability, precision and scoring procedure. *Patient Educ Couns*. 2014;94(1):97–102.
19. Østergaard B, Mahrer-Imhof R, Lauridsen J, Wagner L. Validity and reliability of the Danish version of the 9-item European Heart Failure Self-care Behavior Scale. *Scand J Caring Sci*. 2017;31(2):405–412.
20. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods*. 2016;48(3):936–949.
21. Lin CY, Burri A, Fridlund B, Pakpour AH. Female sexual function mediates the effects of medication adherence on quality of life in people with epilepsy. *Epilepsy Behav*. 2017;67:60–65.
22. Chang CC, Su JA, Tsai CS, Yen CF, Liu JH, Lin CY. Rasch analysis suggested three unidimensional domains for Affiliate Stigma Scale: additional psychometric evaluation. *J Clin Epidemiol*. 2015;68(6):674–683.
23. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Med Care*. 2006;44(11 suppl 3):S78–S94.
24. Brodersen J, Meads D, Kreiner S, Thorsen H, Doward L, McKenna S. Methodological aspects of differential item functioning in the Rasch model. *J Med Econ*. 2007;10:309–324.
25. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care*. 2004;42(1 suppl):I37–I48.
26. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–3191.
27. Harris-Kojetin LD, Fowler FJ Jr, Brown JA, Schnaier JA, Sweeny SF. The use of cognitive testing to develop and evaluate CAHPS 1.0 core survey items. Consumer Assessment of Health Plans Study. *Med Care*. 1999;37(suppl 3):MS10–MS21.

28. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: a conceptual, empirical, and practical guide. *Pract Assess Res Eval*. 2012;17(3):1–13.
29. Cheng C-P, Luh W-M, Yang A-L, Su C-T, Lin C-Y. Agreement of children and parents scores on Chinese version of Pediatric Quality of Life Inventory 4.0: further psychometric development. *Appl Res Qual Life*. 2016;11:891–906.
30. Muthén LK, Muthén BO. *Mplus User's Guide*. 2nd ed. Authors: Los Angeles, CA; 2001.
31. Tsai MC, Strong C, Lin CY. Effects of pubertal timing on deviant behaviors in Taiwan: a longitudinal analysis of 7th- to 12th-grade adolescents. *J Adolesc*. 2015;42:87–97.
32. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measure Transact*. 1994;8:370.
33. Duncan PW, Bode RK, Min Lai S, Perera S. Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil*. 2003;84(7):950–963.
34. Jafari P, Bagheri Z, Ayatollahi SMT, Soltani Z. Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQL™ 4.0 Generic Core Scales in school children. *Health Qual Life Outcomes*. 2012;10:27.
35. Shih C-L, Wang W-C. Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. *Appl Psychol Measur*. 2009;33:184–199.
36. Rosseel Y, Oberski D, Byrnes J, et al. Package 'lavaan'. 2017. https://cran.r-project.org/web/packages/lavaan/lavaan.pdf. Accessed May 31, 2017.
37. Revelle W. Package 'psych'. 2017. https://cran.r-project.org/web/packages/psych/psych.pdf. Accessed May 31, 2017.
38. Andrich D, Hagquist C. Real and artificial differential item functioning. *J Educ Behav Stat*. 2012;37(3):387–416.